ERIC Document Resume	See Instruction Sheet	DATE OF RESUME 8/13/65	DATE RECEIVED

A. PUBLICATION DATA

1. ERIC NUMBER	2. ORIGINATOR CONTROL NO(S).	3. IDENTIFICATION NO.	DO NOT WRITE IN SPACE BELOW
ED000507	7258	N.A.	

4. AUTHOR(S): Fricke, Benno G.

5. TITLE OF DOCUMENT ABSTRACTED: Qualification for Law School and the Bar: How Should it be Determined?

6. PUBLICATION TITLE: N.A.

7. EDITOR(S): N.A.

8. PLACE OF ORIGIN
ORG. University of Michigan
ST.
CITY, STATE Ann Arbor, Michigan
TELE. EXT.
INST/LOC CODE

9. PUBLISHER OR SOURCE: University of Mchigan

10. VOL.	11. PAGES	12. MO./YR.	13. TOTAL PP.
N.A.	N.A.	59	30

MICROFILM COPY AVAILABLE YES ☒ NO ☐

COPYRIGHT GRANTED YES ☐ NO ☒

COPYRIGHT RELEASE FILED YES ☐ NO ☐

TYPE OF REPORT
PROPOSAL ☐
IN-PROGRESS ☐
FINAL ☒
FOLLOW-UP ☐

OTHER ERIC REFERENCES:

ERIC SATELLITE CODE

B. INDEXING DATA

14. SUBJECT AREA

15. TYPE OF ENDEAVOR Report

16. INDEXING TERMS (8-12)

	applicants			
Law education	students			
admissions				

17. ABSTRACT (MAX. 250 WORDS)

A review of findings from research on student assessment and selection for use in appraising applicants for law school and the bar. Presents six recommendations.

ERIC No.

18. SOURCE OF ABSTRACT

C. FUNDING DATA

19. SUPPORTING ORGANIZATION	20. DIVISION OR DEPT.	21. CONTRACT NO. / PROJECT NO.
	22. TOTAL COST OF PROJECT	
	23. DURATION OR PROJECT	U.S. Department of Health, Education, and Welfare Office of Education

ERIC Form 1

IV-19 (Law) Complete ref.

This has been published in the volume Law Schools Look Ahead but I have no reprints. BGF.

Qualification for Law School and the Bar: How Should it be Determined?

Benno G. Fricke

Assistant Chief, Evaluation and Examinations Division and Assistant Professor of Psychology, University of Michigan

Shortly after I was asked to prepare a paper on "student assessment and selection" for leaders in legal education I began a search for pertinent literature in the library of the University of Michigan Law School. A few minutes in the library (where I learned, for example, that about 75 law schools publish law reviews) convinced me of several things, most important of which was the desirability of brevity. Accordingly, in this paper I have attempted to present what might be called a relatively short summary of important points and views from assessment psychology which might be helpful to busy deans and other legal educators as they go about their work shaping the destiny of the law, and indirectly, but in a real sense since so many governmental officials are lawyers, the future of the nation and the world.

The paper's title suggests that some answers might be offered for the solution of some long standing problems. To increase the probability that my suggestions will be taken at least somewhat seriously it is best that they not be presented until after some background information has been shared with you.

At the outset it should be recognized that my objective is not to disclose a means whereby the law profession can corral for itself the top talent of the nation. As an assessment psychologist concerned with the optimum use of the nation's human resources I think this would be a mistake, but a mistake which would probably result in the prevention of governmental blunders such as now are being witnessed in Michigan! Just how much of the nation's top talent should be taken or accepted by any one profession

ED000507

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education

This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

or discipline is a difficult question to answer, and this is not an appropriate place to attempt it. The major suggestions in this paper are intended to bring about a rearrangement and improved use of the talent currently claimed by the law schools and the law profession, and to bring about tapping the portion of the total supply of top talent which is for all practical purposes currently unused.

I have selected four important findings from student assessment research that you should know about and take into account in appraising applicants for law school and the bar. The first of these is that judgments made by human beings are usually less accurate than judgments made mechanically from a statistical table or formula. Of course this sounds unbelievable, particularly to admissions officers who feel they really understand applicants after they have considered a great deal of information about them (e.g., aptitude test scores, previous academic grades, letters of recommendation, interview impressions, etc.). It is true that I have no data to show how good an assessment job the admissions officers of law schools can do, but I am confident a research study would show actuarial forecasts based on only two items they routinely use (previous grades and an aptitude test) would provide predictions which are slightly superior. Almost everyone unfamiliar with the research results, some of it ably reviewed in a book by Paul E. Meehl,[1] is convinced that if he were put to the test, the mechanical approach which can be handled by a competent clerk, would be found inferior.

Several years ago in an article[2] addressed to the nation's admissions officers

---

[1] Meehl, Paul E. Clinical Versus Statistical Prediction: A Theoretical Analysis and a Review of the Evidence. Minneapolis: University of Minnesota Press, 1954.

[2] Fricke, Benno G. Prediction, Selection, Mortality, and Quality Control. College and University: Journal of the American Association of Collegiate Registrars and Admissions Officers, 1956, 32, 34-52.

I urged, among other things, that each admissions officer see how good a judge he was, and specifically to see if his appraisal of an applicant's academic promise was superior to a formula. So far as I know, no officer has done this.

Since that time I have analyzed pertinent data from two large samples of the nation's leading admissions officers—all officers for colleges having membership in the College Entrance Examination Board. The results from the comparisons, along with additional material from a questionnaire study of the opinions of 300 officers, will come as no surprise: the officers did not agree with each other in their evaluations; their assessments were inferior to those made by a formula; and they do not wish to admit students by means of a formula.[3] The admissions officers are right when they say the formula is fallible. But they, and probably all other judges of student quality, are unable to indicate in advance when the formula will be in error. The important question is not what method is infallible but rather what method is best. Errors will always be made. By every method some applicants who appear to have high promise and are enrolled will perform unsatisfactorily. And by every method some applicants who appear to have inadequate potential and are rejected would have performed satisfactorily if they had been admitted. The objective should be to keep the number and extent of errors to a minimum.

You might be interested in the specific results and consequences of a related study. It is unique in that actual admissions decisions were made on the basis of two different selection procedures. The object of special interest was the value and cost of the widely used and highly regarded "selection interview." The study was conducted in Harvard University's Graduate School of Business Administration.

---

[3] Fricke, Benno G. Tests, Students, Standards and Admissions Officers. Long-Range Planning for Education (Traxler, Arthur E., Editor) Washington: American Council on Education, 1958. Also reprinted as How Colleges Should Pick Their Students. College Board Review 1958, 34, 17-22.

Vernon R. Alden, Associate Dean of the School, gave an account of their experience.[4] Here is what he had to say:

"When the Harvard Business School discovered that a substantial portion of its $61,000 admissions budget was being spent on personal interviews with every applicant for admission, an evaluation of this time-consuming procedure was launched. We discovered that some members of our admissions board, several members of the faculty, many alumni, and administrators in other colleges believed that the evaluations made of students in the personal interviews were often faulty and sometimes even harmful. Consequently, we undertook a study of the value of personal interviews in actual admissions decisions for the class entering in the fall of 1952.

All 2,000 men who applied for admission to the 1952 class had been interviewed in accordance with our regular procedure. As applications were received, they were processed through two different admissions boards—but only one of these committees was given the interview data. Both, of course, had all the rest of the data before them. When the final results were in, we found that the two boards had agreed on 96 per cent of the candidates.. To carry the experiment one step further, we took the 80 men on which the boards had parted company and split it evenly taking 40 admitted by one group and 40 by the other. As you might suspect we studied both groups very carefully during their two years at the school. The 40 men admitted with strong college records and poor interviews clearly did a better academic job. Moreover, when faculty members and fellow students were queried about other personal qualities of the men in both groups, those with the strong papers and poor interviews again showed up better than the strong interview group.

Two years later we acted on the basis of this experience and did away with the compulsory interview requirement on an experimental basis. We are quite willing to talk with students who wish to visit and ask questions about the Business School, but we keep no records. This second experiment is in its third year, and so far we have seen no difference in the classes which we have been admitting. If anything, our faculty members tell us that these recent classes have been doing a better job at the school. In the dean's office we have noticed no increase in the number of misfits, no lack of participation in extracurricular and social activities, and no evidence of lesss school spirit. Our company interviewers, who would be, no doubt, the first to notice a change in the character of our graduating students, have been enthusiastic about the men in the last two Master of Business Administration classes.

---

[4] Alden, Vernon R. College Finances in Today's Economy. College Admissions: Volume 5. New York: College Entrance Examination Board, 1958.

We intend to continue our experiments with admissions interviews, perhaps modifying our present position to some extent. But one thing is already clear to us—requiring personal interviews is expensive, both to the school and to the applicants for admission. We have come to the conclusion that we need clear evidence, which we have been unable to find, to justify reinstating the interview requirement as a part of our admissions program. To me, this experience has been an eye-opener. From my experience working in two other university admissions offices—at the undergraduate level—plus my activities at the Harvard Business School, I firmly believed that an admissions office could not select a high-quality class without interviews. I am now ready to admit that my subjective view on interviews could not be supported by evidence. As a result of this experience, I feel constrained to look at other "sacred cows" in our various operations at the school to see whether they, too, can be justified in terms of costs."

Dean Alden need not look far. While it is reassuring to find that he is convinced that the selection interview serves to lower rather than to raise the quality of admitted students and is therefore an unnecessary cost, a further increase in student quality and a further reduction in unnecessary cost would result by <u>complete</u> removal of human judgment, and dependence on a statistical formula employed mechanically by a clerk. It is important to note that all who were admitted, including the two groups of 40 students, were admitted on the basis of the <u>judgment</u> of an admissions committee, and all whose applications for admission were rejected were rejected on the basis of human judgment. One can be confident that some who were admitted were inferior to some who were rejected; fewer errors would probably have been made if the decision to admit or to reject were made without the use of high-cost human judges.

Incidentally, anyone interested in a thorough treatment of the selection interview and its usefulness (or, more accurately, its uselessness) can find it in excellent articles by E. Lowell Kelly[5] and Lewis B. Ward.[6] They review not only

---

[5] Kelly, E. Lowell. A Critique of the Interview. <u>Journal of Medical Education</u>, 1957, 32, Part 2 of No. 10, 78-84.

[6] Ward, Lewis B. The Interview as an Assessment Technique. <u>College Admissions: Volume 2</u>. New York: College Entrance Examination Board, 1955.

opinions on the interview but also the research results. For our purposes it will suffice to say that the selection interview is an expensive item higher education in this country cannot afford, especially since it seems clear that it is a major factor in locating students at inappropriate institutions. Good students who do not look good to the interviewer too often are rejected, and poor students who do, are admitted. While it is not known how many of the nation's students find themselves in inappropriate institutions, the number is certainly much too large, especially at a time when optimum use of the nation's human resources should be made. Those who use their judgment when something better is available are not rendering a service even though they believe strongly that they are.

The second of the important findings disclosed by assessment research which have implications for those concerned with legal education and standards is the great diversity in student and in institutional quality. That these differences exist will surprise no one connected with higher education, but the extent of the differences is not well recognized, and if recognized it is not adequately taken into account.

In a fact-filled book[7] Dael Wolfle presented data on the aptitude (IQ's) of freshmen who entered 41 colleges in 1946. The freshmen enrolled in the college having the lowest average score, 108, were not uniformly weak. Half the freshmen made scores between 99 and 117. The bottom quarter had scores below 99 and the top quarter had scores above 117. Similarly all freshmen in Wolfle's top-ranking college (average 131) were not uniformly strong. One quarter had scores below 126 and one quarter had scores above 137. Wolfle indicated that both colleges at the extremes were private, nonsectarian, and accredited.

---

[7] Wolfle, Dael. America's Resources of Specialized Talent. New York: Harper Brothers, 1954.

Traxler[8] examined the aptitude scores of freshmen in the 323 colleges, mostly publicly supported ones, comprising the norm group for the 1937 edition of the American Council on Education Psychological Examination (abbreviated ACE). By converting the ACE scores to IQ's he estimated the range in average IQ was from 94 to 123. Since the average IQ of high school graduates is about 105 it is clear that the standards in some of the 323 colleges must be rather low. The range in IQ would have been much greater than 29 points if the scores of freshmen in all 1800 colleges in the country were taken into account.

Data I have collected from admissions officers and other sources supports the view that differences among students enrolled in any one college, and differences among institutions, both public and private, are tremendous. For example, at the University of Michigan approximately 2% of the entering freshmen have IQ's below 100 and approximately 25% have IQ's above 130. From statistics on the 203 colleges in the 1933 norm group[9] for the ACE it can be estimated, for example, that less than 1% of the freshmen enrolled at the six top ranking colleges obtained scores below the average score obtained by freshmen enrolled at the six bottom ranking colleges; similarly less than 1% of the freshmen at the bottom six colleges obtained scores above the average freshman at the top six colleges. These "percentage overlap" estimates are made by means of Tilton's tables.[10] It is fairly certain that if an average or

---

[8] Traxler, Arthur E. What is a Satisfactory IQ for Admission to College? School and Society, 1940, 51, 162-164.

[9] Thurstone, Louis L., and Thurstone, Thelma G. The 1933 Psychological Examination. Educational Record, 1934, 15, 161-175

[10] Tilton, J. W. The Measurement of Overlapping. Journal of Educational Psychology, 1937, 28, 656-662.

typical freshman from one of the top colleges enrolled in one of the bottom colleges he would achieve a very superior record; similarly if an average freshman from one of the bottom colleges enrolled in one of the top colleges he would fail.

Institutional differences in aptitude take on special importance when the college grade records of applicants from different institutions are to be evaluated and compared. Is a grade-point average of 2.89 at X, a below average difficulty institution, equivalent to a grade-point average of 2.07 at Y, an above average difficulty institution? Correct answers to questions such as this are difficult to arrive at even when much is known about the grading practices of X and Y. According to data from a variety of sources, as well as that supplied by the 300 admissions officers referred to above, it seems that approximately the same grades are given at institutions enrolling superior students as are given at other institutions. Typically about 25% of each college freshman class receive unsatisfactory grade-point averages. For example, the percentages of freshmen given unsatisfactory grades at Amherst, MIT, and Swarthmore is 21 - 30 according to the admissions officers, the same percentage of Austin Peay State College, Beaver College, and University of Kentucky freshmen receive such grades. At California Institute of Technology, Dartmouth, and Harvard the percentage unsatisfactory is 11 - 20; the same percentage is found at Arizona State College, Casper Junior College, and Villanova. This provides confirmation of the view that it is probably incorrect to describe an individual student as "a B student," or "a C plus student," and so on. A freshman with a C+ average at a particular institution might very well have earned an A average at another institution where somewhat inferior students are enrolled or a D average at another institution where somewhat superior students are enrolled. A much more accurate description of an individual student would be "he's a Harvard B," or "he's an Indiana B," or "he's a Bay City B," etc.

The data also support the point that there seem to be no absolute standards of satisfactory performance. For example an examination paper written in a particular course at a particular institution might be evaluated by the instructor as being of C (or satisfactory) quality; the same paper evaluated by an instructor from a different institution accustomed to examination papers written by inferior students might evaluate it as being of A (or excellent) quality; but the same paper evaluated by an instructor from another institution accustomed to examination papers by superior students might evaluate it as being of D or E (or unsatisfactory) quality

This is perhaps an appropriate place to mention some of the results from a fairly old and excellent study by Learned and Wood[11] of the achievement of college seniors and high school seniors. Learned and Wood administered a lengthy achievement examination consisting of questions in (a)history and social studies, (b)natural science, (c)fine art and world literature to students in the high schools and colleges of Pennsylvania. You have probably guessed that some of the high school seniors knew more than some of the college seniors. But you might be surprised to know that about 10% of the high school seniors got higher scores (i e , knew more) than the average college senior. And, perhaps even more depressing, about 10% of the college seniors got lower scores than the average high school senior. Certainly it seems fair to say that a Bachelor's degree is not awarded on the basis of how much one knows, for if it were many high school seniors, who in terms of "educational time" are not quite ready to begin college, should be given a degree and many college seniors who are almost finished should not be given a degree. (Perhaps they should be given instead a second high school diploma!) Obviously it is incorrect to assume that because a person has obtained a college degree he is well educated; all that can be assumed is

---

[11] Learned, W. S. and Wood, Ben D. The Student and His Knowledge. New York: Carnegie Foundation for the Advancement of Education, 1938.

that he spent four years in college. In many respects higher education institutions are like penal institutions--release is dependent not upon demonstrated behavior but upon time spent.

Before leaving the topic of quality differences in students and in institutions, the heterogeneity of ability in, and among professional schools merits comment. Many who believe students in high school and college are rather heterogeneous believe students in professional schools and graduate programs are not. Are all law students equally able? Reference is made here to a study reported by Marjorie Olsen[12] on the predictive value of the Law School Admission Test (the LSAT provided slightly though not significantly superior predictions of success than pre-law record in most schools), but our present interest lies not in correlations but in the average scores and the variability of scores for different schools. Of the 20 schools on which she reported I selected two which were characterized as "large," which differed most in quality, and which were involved in both the 1948 and 1951 studies.

Table 1 shows the relevant statistics. It is obvious that the difference in LSAT scores between these two schools is very great. According to Tilton's tables less than 3% of the students in school 5 surpass the average student in school 9, and less than 3% of the students in school 9 fall below the average student in school 5. It is highly unlikely that these schools can make similar intellectual demands of their students even though their structure, stated purpose, and curriculum may be similar. An LLB from one does not mean the same thing as an LLB from the other, and the public cannot and should not be led to expect the same things from those who possess them.

---

[12] Olsen, Marjorie. The Law School Admission Test as a Predictor of Law School Grades, 1948-1953. Princeton: Educational Testing Service, 1955.

Table 1

LSAT Percentage Overlaps, Means, and Standard Deviations, and Correlations with Freshman Grades in Two Law Schools

Student Group	Percentage Overlap	LSAT		Correlation with Law School GPA	
		Mean	SD	LSAT	Pre-L GPA
School 5 - 1948	2.1	437	90	.52	—
School 9 - 1948		589	60	.44	.41
School 5 - 1951	2.7	431	97	.48	.36
School 9 - 1951		588	65	.46	.43

It would be unfair to leave the impression that great diversity is not characteristic of other leading professions. Since the situation in medicine is of continuing interest to those in law the following information may be worth noting. Professor Thomas R. McConnell, Director of the University of California's Center for the Study of Higher Education, in a paper[13] presented at the 1959 Annual Meeting of the American Association of Collegiate Registrars and Admissions Officers disclosed that "The equivalent average ACE score of seniors in the medical schools in 1956 was about 125; the mean for college freshmen in our national sample was approximately 104. The school means of medical seniors ranged from 107 to 159 in ACE scores, with the lowest only slightly above the mean of undergraduate college entrants. And 16 per cent of the medical seniors had equivalent ACE scores below the mean of

---

[13] Professor McConnell's paper "Do Students Make the College?" will probably appear in the summer issue of College and University.

college freshmen in general. At the other extreme, four schools appear to have unusually large proportions of the academically capable, with class means above 150." In commenting further on the diversity in medical students he stated "When one realizes that almost 25 per cent of the medical seniors have academic ability scores no higher than the average college graduate and about 25 per cent at the other end of the continuum have scores equivalent to or higher than the average person completing a Ph. D. degree, one must conclude that the character and quality of programs varies considerably."

A third finding which has important implications for those concerned with the evaluation of college students lies in the stability of the quality of students in different institutions. It is simply not true that an institution enrolling superior students this year will enroll inferior ones next year. Minor changes occur of course. However, the differences in institutions discussed above are sufficiently consistent over time to permit fairly accurate inferences to be drawn from pre-law records established at different institutions at different times.

To illustrate institutional stability over time I have assembled the material in Table 2. It shows the ACE scores of freshmen enrolled in 38 colleges in 1933 and 1952. The 38 colleges were among the 203 colleges in the 1933 norm group[14] and they were also among the 259 colleges in the 1952 norm group.[15] Thus, all colleges common to both norm groups are involved. In both norm groups many colleges had higher and many had lower scores than the 38 ranked in Table 2. More specifically, Trinity College and Eastern Kentucky State College were not among the top and

---

[14] Thurstone, Louis L., and Thurstone, Thelma G. Op cit 9.

[15] Cooperative Test Division. Norms Bulletin Princeton Educational Testing Service, 1953.

Table 2

Comparison of Average Academic Aptitude Test Scores (ACE's)
Obtained in 1933 and 1952 for Freshmen in 18 Institutions

Institution	1933 Freshmen		1952 Freshmen	
	Median	Rank	Mean	Rank
Trinity, Hartford, Conn	213.6	1	122.3	4
Wells, Aurora, N. Y.	199.0	2	120.7	5
Clark, Worcester, Mass.	189.2	3	117.0	8
Wheaton, Wheaton, Ill.	183.3	4	116.3	9
St. Xavier, Chicago, Ill.	181.0	5	115.6	10
Tufts, Medford, Mass.	180.0	6	123.6	1
Colleges of the Seneca (Hobart and William Smith) Geneva, N. Y.	177.5	7	111.7	16
College of St. Elizabeth, Convent Station, N. J.	175.1	8	108.4	20
Lawrence, Appleton, Wis.	175.0	9	113.6	11
La Salle, Philadelphia, Pa.	174.1	10	107.5	22
Kalamazoo, Kalamazoo, Mich.	173.3	11	119.5	6
De Pauw, Greencastle, Ind.	171.4	12	113.0	12
Alfred U., Alfred, N. Y.	164.5	13	113.1	13
Union (Iff), Sioux City, Ia.	162.5	14	105.5	24
[illegible] Vermont, Darlington, Vt.	161.2	15	111.4	14
Bethany, Bethany, W. Va.	160.9	16	107.3	23
Hood, Frederick, Md.	160.8	17	112.6	13
Parsons, Fairfield, Iowa	160.1	18	98.6	32
Immaculata, Immaculata, Pa.	160.0	19	109.6	17

over

Bradley, Peoria, Ill	156.3	20	96.3	34
Rosemont, Rosemont, Pa.	154.4	21	120.2	4
Park, Parkville, Mo	153.9	22	105.7	25
Goshen, Goshen, Ind.	153.0	23	100.9	29
St. Vincent, Latrobe, Pa.	151.4	24	105.9	24
Gettysburg, Gettysburg, Pa.	145.6	25	112.8	12
Washington and Jefferson, Washington, Pa.	143.9	26	114.9	10
Washington, Chestertown, Md.	143.3	27	108.8	18
Franklin Coll. of Indiana, Franklin, Ind.	142.0	28	105.3	26
Hanover, Hanover, Ind.	141.9	29	105.1	27
Loretto Heights, Denver, Colo.	141.7	30	92.3	36
St. Mary of the Woods, St. Mary of the Woods, Ind.	138.3	31	107.9	21
Central, Fayette, Mo.	135.2	32	100.7	30
W. Va. U., Morgantown, W. Va.	133.6	33	99.1	31
Simpson, Indianola, Iowa.	124.2	34	105.0	28
Alabama, Montevallo, Ala.	122.1	35	94.2	35
Tusculum, Greeneville, Tenn.	121.8	36	87.8	38
Spring Hill Coll., Mobile, Ala.	115.0	37	97.8	33
E. Kentucky State College, Richmond, Ky.	103.5	38	89.2	37

bottom six colleges referred to above. (Incidentally the percentage overlap between Trinity and Eastern Kentucky is less than 3%, the same overlap found between the two law schools. That is, the difference in the aptitudes for the two sets of institutions is identical--Trinity, is as much superior to Eastern Kentucky as Law School 9 is to Law School 5.)

Despite the fact that the test was shortened and its content altered over the 19 year span an amazing stability in the rankings is to be noted. The rank order correlation coefficient is .78, which is highly significant.

A fourth general area of research on quality and promise of students has to do with the predictive and assessment value of aptitude tests and particularly the LSAT. A rather thorough review of studies and tests up to 1951 was reported by Feeney.[16]

Perhaps the most remarkable finding is the <u>similarity</u> in predictive coefficients. Essentially the same validity coefficients are found for all schools, and for all tests designed to predict school or college achievement. Almost always the coefficients turn out to be about .45, rarely are they less than .30 or more than .60. The size of this coefficient disturbs me even more than the tremendous range in the average ability of students in different institutions. You may be surprised to know the reason why I am disturbed by the size and stability of the coefficients--they are <u>too high</u>. High predictive coefficients (and they tend to be almost .60 when LSAT and pre-law grades are combined into a single index) I believe are undesirable and should be reduced to zero if possible. They reflect inequalities of opportunity for success.

I am aware that this view concerning high coefficients is atypical. The

---

[16] Feeney, Bernard J. How Good are Legal Aptitude Tests? <u>Journal of Legal Education</u>, 1951, 4, 69-85.

common conception is to be found in the report of Burnham and Crawford,[17] for example. In describing their study at Yale, where the mean LSAT was 599 and the SD 65, they state "the correlation is only .37" between LSAT and low grades. After the pre-law record was combined with the LSAT a higher coefficient was obtained. Their comment on it is, "Again, the correlation (r = .53) is less than desirable, but it is still high enough to be of considerable use."

Statements such as this confuse the values and objectives of test researchers and of test users. Researchers tend to be interested primarily in prediction of performance; users are, or should be, more interested in the assessment of quality. Perhaps the situation can be clarified through an extreme hypothetical case. Suppose, for example, that you had the choice of all prospective law students in the country. Let us suppose that you were aware that the LSAT tends to be the best predictor of success and that you decided to select for your school the 125 students whose LSAT scores were 732, 733, and 734. The mean would be about 733 and the SD almost 00. At the end of the year you correlate the LSAT scores of your 125 students (or the number who finish the year--some will prove to be too weak to handle the program designed and appropriate for the group) with their freshman grade-point average. The correlation would almost certainly be less than .01. Now how do you interpret your results, and what actions would you take?

You might conclude (as many have unfortunately, on the basis of somewhat similar data) that the LSAT is worthless, and that if you had first choice again you would not even look at the LSAT scores of applicants. You might regard the coefficient as undesirably low--the high scorers (those with scores of 784) do not obtain better records than the low scorers (those with scores of 782). You might feel that any

---

17 Burnham, Paul S. and Crawford, Albert B. Law School Prediction at Mid Century. Journal of Legal Education, 1957, 10, 189-200.

rationally acceptable item such as number of books read last summer, which for example correlated .16 with grades, would be a better basis for selecting successful students.

But you might also make the following interpretation. The validity coefficient of .01 and the mean of 78 reflect a desirable state of affairs. The students are uniformly strong and all have an equal opportunity for success. None are destined, i.e., to have a high probability of failure or mediocrity. The grades and class rankings are not merely a reflection of general aptitude. All have enough of that. Factors other than aptitude (e.g., how a student applies himself) determine success and failure.

Schools should not want to be able to predict the performance of admitted students. If high coefficients are desired they may be obtained by simply admitting all applicants. It would then be a simple thing to predict before classes begin, those students who are likely to fail and those who are likely to succeed. There are few who would argue that this is the way to admit students. If it can be said with considerable accuracy at the time a student enters a particular institution that he will fail, he should not be admitted. A concrete way of seeing the significance of the relationships now found in most law schools is to examine the law grades of those who are at the extremes on an index based on a combination of LSAT and pre-law GPA. In general, those who rank in the bottom 5% of their freshman class have about one chance in twelve of doing as well as or better than the average student. Those who rank in the top 5% of their class have about eleven chances in twelve of doing as well as or better than the average student.

Although the concern of educators is usually with the small but weak minority and the probable difficulty they will have doing satisfactory work, there are reasons why it is undesirable for the small but strong minority to have success assured them (e.g., it would be better for society and these strong students if they were to compete

only with other strong students.) From the standpoint of one concerned with the optimum use of human resources it would be best if applicants who are far below or far above an institution's average were refused admission and advised to go to an institution where they are about average.

Obviously my disagreement with Burnham and Crawford's view, and my preference for low predictive coefficients should not be interpreted as favoring the use of invalid selection devices (such as size of ear lobe) because they provide low or zero correlations. It seems to me there are at least two steps or phases in prediction and assessment which should be distinguished. The first is the search for or creation of instruments, factors, devices, methods, etc., which are valid predictors of academic success. (At the present time my colleagues and I are developing personality and motivation scales to predict academic success -- and we are pleased with the relatively high predictive coefficients obtained so far, we are doing everything possible to increase them.) The second phase is the use of these instruments, factors, devices, etc., in such a way that they do not reduce their predictive validity within a selected group of students. (After we are certain that the personality grade-predictor scales have adequate predictive validity they will be turned over to admissions officers, counselors, and others. These test users should reject applicants who have low scores because their prognosis is poor, and advise low scoring counselees that they should spend more time with their books and less on extracurricular activities than the typical student if they expect to achieve average grades.)

Perhaps it should also be pointed out that to obtain a zero or low correlation for the LSAT it is not necessary that everyone selected have the same or almost identical score. Considerable variability in LSAT scores may be present. It is possible for such variability and low coefficients to result from enrolling students who have high LSAT's and low pre-law GPA's (pre-law GPA, it will be recalled, also

correlates substantially with law school success), and students who have low LSAT's and high pre-law GPA's. For example, two students having a very high LSAT (say 790) and a very low pre-law GPA (say 1.04) cannot perform in line with both these indicators. If one student achieves a very good record and the other student a very poor one at the same law school the correlations for each predictor variable will be zero. That is, a very low pre-law GPA is not necessarily a sure sign of failure. If it is coupled with a very high LSAT anything can happen, and this means zero predictive accuracy.

Too frequently the two phases, the "in research" phase, and the "in use" phase are confused. Correlational studies in the first phase permit an evaluation of the instrument, factor, etc.; correlational studies in the second phase permit an evaluation of the use of the instrument, factor, etc.

It is clear from the coefficients reported for the LSAT that greater weight should be given the LSAT in the admissions process— all of the "valid variance" has not yet been exhausted.

In light of the above findings and concepts from assessment and prediction research, what changes, if any, might be made with respect to the selection and education of students for law? Joiner[18] has recently pointed out the importance of planning alternative ways of handling the expected large number of law school applicants. Whether the law will attract relatively more (and better?) students in the future is not clear. Legal educators predict a large increase in the number of positions for lawyers. These predictions made today may be right— but they are the same ones that were incorrectly made many years ago. Great changes in the public's perception of lawyers and their services are not likely to come about unless rather specific things are done to bring them about. Are there too many lawyers at the present time? No answer seems to be agreed upon.

Joiner, Charles W. The Coming Deluge: How Goes Our Ark? Journal of Legal Education, 1957, 9, 466-472.

Reginald Heber Smith[19] has presented figures to show that the number of admissions to the bar per population million has been decreasing for at least the past 25 years. This does not necessarily mean there are too few lawyers to meet society's felt needs. From all I have been able to gather there seem to be too many with the label "lawyer," and a shortage of "able and specialized lawyers." To a layman the distinctions among lawyers are not very obvious. The differences in interests, abilities, and training of lawyers probably should be made more apparent than they are today. There seem to be too many "GP's" who are willing to tackle almost any assignment. An interesting professional account of the pros and cons of specialization in the law appeared several years ago.[20]

Perhaps law schools might increase the development of specialty education and gradually work toward differentiation of lawyers. It would seem best for schools with very superior student bodies to prepare lawyers for the more intellectually demanding branches or specialties. Schools with relatively inferior student bodies might prepare lawyers for the less demanding branches or for general "small town law." Large law schools might be able to offer a considerable number of specialties. Each law school probably should continue to offer what might be called a core or basic

---

[19] Smith, Reginald Heber. The Bar is Not Overcrowded: Some Facts About an Ancient Legend. American Bar Association Journal, 1958, 44, 1054 +.

[20] Joiner, Charles W. Specialization in the Law: Control it or it Will Control You. American Bar Association Journal, 1955, 41, 1105 +.

Wham, Benjamin. The Public Need Must Be Better Served. American Bar Association Journal, 1956, 42, 39+.

Siddall, Roger B. Retort to Professor Joiner's Call for Control. American Bar Association Journal, 1956, 42, 625+.

program. If different law schools continue to enroll students who differ markedly in quality it would seem wise to recognize the diversity and to plan appropriate educational programs for them. By doing something like this the employers of lawyers and those who avail themselves of their services will know more precisely what they may expect.

It follows from what was said above about the undesirability of high predictive coefficients that each law school should select students from a relatively narrow range of academic quality. This would reduce the coefficients and the student mortality rate and permit an educational program to be conducted at a "challenging-to-all" level. A school enrolling students having a mean LSAT of 430 and a mean pre-law GPA of 2.35 is unable to challenge adequately an applicant with an LSAT of 650 and a GPA of 3.70 at a leading independent college. Of course he would be a "success" as far as law grades are concerned but he would probably grow much more if enrolled at an institution enrolling superior students. I believe this school should refuse to admit such a student and refer him to a more appropriate one.

The phrase "qualified for law school" as commonly used lacks meaning. An applicant may be qualified for one school but not for another, and he may be over-qualified for still another. In my opinion a student is qualified for a school if the achievement you predict for him is approximately at the class average; if his predicted achievement is below average he is not qualified, and similarly if his predicted achievement is above average he is overqualified (or too able). If he is predicted to be below average or above average his needs can probably best be met by some other school.

Now, just how should the qualifications of applicants be determined? Certainly much local work needs to be done to arrive at the best answer for each school. Some general suggestions however might be worthwhile. The first one is to

reduce the evaluation and decision process to a formula of some kind, and eliminate or drastically reduce the role of the selection interview and letters of recommendation. Elaboration of this point seems unnecessary. The second suggestion is to publicize the formula (or formulae) you use to arrive at decisions. This probably would attract many desirable applicants who otherwise would not submit applications (e.g., because they heard "it's impossible to get into X unless you have connections," etc.); it would also discourage applicants who are obviously not sufficiently able.

The third suggestion is to establish more than one cut-off point. A single cut-off point above which all are qualified for admission and below which all are unqualified would be unwise unless the number to be accommodated can be very flexible. Possibly two cut-off points might be enough for the highly selective law school desiring to get the most able students it can. One point might be set so high that only, say, 25% of currently enrolled freshmen at the school would qualify. The law school catalog would state "any applicant with a qualification index of X or higher will be admitted automatically without further consideration." The second cut-off point might be set so low that only, say, 15% of enrolled freshmen would be rejected automatically. The catalog would state "a student with an index of Y or lower is not sufficiently strong to merit consideration for admission." The catalog would also state that applicants not falling into these extreme groups of "clearly admissible" and "clearly inadmissible" are encouraged to submit the regular complete application for admission and that they would be notified as soon as a decision was made.

The fourth suggestion, for those who want a very rough formula to depart from, is to give approximately equal weight to the LSAT and pre-law GPA (since research shows that both have approximately equal predictive validity.) This might be done by adding the LSAT score to a weighted pre-law GPA score and dividing this sum by two to give the Law Qualification Index (LQI). Until local research devises

a more accurately weighted pre-law score it is suggested that the "sending colleges" be divided into only three groups: above average difficulty colleges (the AA's), average difficulty colleges (the A's), and below average difficulty colleges (the BA's). The AA's would include a large majority of the non-sectarian private (or independent) colleges comprising the College Entrance Examination Board, and a few other independent colleges not in the CEEB. The A's would include almost all the rest of the CEEB colleges, the major state supported universities, and a considerable number of independent colleges not in the CEEB. The BA's would include the rest (e.g., all but one or two of each state's publicly supported institutions, most of the private colleges that are struggling for survival, etc.)

In what follows it is assumed that the pre-law GPA is expressed in the "A has 4 honor points, B has 3" form or can be readily converted to it.

To transform the pre-law GPA established at the AA, A, and BA colleges into roughly comparable achievement units the following "college multipliers" are suggested: for AA colleges 250, for A colleges 200, and for BA colleges 150. The product of the college multipliers and the pre-law GPA is intended to produce a distribution of scores with an approximate mean of 500 and a standard deviation of 100 so the units would be roughly comparable to those on the LSAT. It is of course highly unlikely that the values are optimum; a systematic study, preferably national, in cooperation with graduate and other professional schools, should be made which would assign a specific multiplier to each undergraduate college or small group of highly similar ones.

The Law Qualification Index (LQI) of a student with a pre-law GPA of 2.60 at Haverford College and a LSAT of 610 would be $\frac{(2.60 \times 250) + 610}{2} = 630$

The LQI of a student with a pre-law GPA of 3.40 from Western Michigan University and a LSAT of 530 would be $\frac{(3.40 \times 150) + 530}{2} = 520$

A certain law school, L, having stated it would admit applicants with LQI's of 600 and above, refuse applicants with LQI's of 500 and below, and consider applicants with indexes below 600 and above 500 would send an admission certificate to the Haverford student and a "cannot give a definite answer yet" letter to the Western Michigan student.

Thus the college multipliers take into account the major differences among institutions, and pre-law GPA takes into account the achievement differences among students within these institutions. The LSAT score in the formula takes into account the aptitude differences among the students.

It is not enough to know a student earned a 2.60 at Haverford. If the applicant above had had a LSAT of 340 rather than 610 it is unlikely that he would be able to meet the standards at school L; since his LQI would have been 495, admission would have been refused automatically.

As indicated above from the standpoint of optimum use of the nation's legal resources it would be wise for each law school to indicate the quality of students for which it has an appropriate program, and whom it will admit. Law School M, unlike L which has a relatively superior student body, draws primarily average ability applicants. It s catalog might state it "will admit all who have LQI's between 470 and 530, will consider all who have LQI's of 440 to 470, and 530 to 560, but will refuse all who have LQI's below 440 and above 560." Law School O which has a relatively inferior student body might state it "will admit college graduates with LQI's below 400, will consider applicants with LQI's of 400 to 440, but will refuse applicants with LQI's above 440.

The fifth suggestion is to admit certain promising high school graduates directly to law school (or to certain law schools.) Perhaps it would be desirable to

arrange a special program of five years, rather than the typical six or seven year ones, for eager and able high school graduates. It is my firm conviction that if interested high school students in the top 10% of their graduating class who also have scores above 600 on the CEEB's Scholastic Aptitude Test (SAT) were enrolled in a five year law program today they would, on the average, prove to be much superior to the average regularly trained and graduated lawyer in seven years. The findings of Learned and Wood should be recalled in this connection. One should be aware that mere possession of a college degree is no guarantee of high quality; yet it is apparent that admission to many law schools is more or less assured by it. While the whole question of time spent in college versus amount of educational growth and promise is relevant, a satisfactory discussion of it would be too space consuming. However, one interesting incident described by E. Lowell Kelly[21] should be passed on.

"In the study we are doing at the University of Michigan we find the best single predictor of general medical achievement (as measured by over-all grades and National Board Examination scores) to be premedical college grades. We also find, curiously enough, that the number of premedical credit hours in inorganic chemistry and in biology submitted by the applicant is negatively correlated with medical school grades. This is not surprising, I think. The weak student, who feels himself weak, or whose advisor feels the student is weak, is very apt to take additional courses in chemistry or biology. Medical educators may not like it, but at least some misguided advisers recommend this. At any rate, we find a significantly negative correlation between the number of hours in such courses and our criterion of medical achievement."

While I suspect that the very promising high school graduates could compete favorably with the college trained law freshmen it might be best to provide separate courses, or at least sections for them. It would seem best if during the first two years (of the five) only about half the course work was legal and the rest "pre-legal." Perhaps during the last two years students might spend half their time on some kind of practicum or internship program. Well advertised legal aid clinics established within commuting distance of the school would not only give students experience, but

---

[21] Kelly, E. Lowell. Op. cit. 5.

it would also provide an opportunity for the law to inform the public of possible legal services; a small service charge for those who can afford it might contribute toward paying the salary of the supervising instructor.

In some respects one of the greatest benefits from admitting high school graduates directly into a five year program would be the increase in the percentage of women in law school. Many capable women are now not able to spend seven years to get an LLB. Probably one of the best relatively untapped sources of gifted individuals in this country (unlike in Russia) is women. In many ways they are ideally suited to the study and practice of law—their verbal and reading abilities tend to be higher than those of men.

Many able young men not interested in the mathematically oriented profession of engineering would probably be influenced to enroll in law directly after high school if this were possible. It is likely that many of these students in the past have found other non-mathematical fields which interest them during their two to four years of college.

Before leaving the high school and pre-legal education topics a comment or two concerning a possible conflict in the standards suggested by the above and by the Association of American Law Schools may be in order. The Committee on Pre-Legal Education,[22] while not prescribing any particular courses felt all students entering law schools should have adequate "A. Comprehension and expression in words; B. Critical understanding of the human institutions and values with which the law deals; and C. Creative power in thinking." It is apparent from the Committee's description of these items that the LSAT, the College Board's SAT, and most intelligence-academic aptitude tests give quite good measures of items A and C. I know of no better measures.

---

[22] Committee on Pre-Legal Education. On the Aims and Quality of Pre-Legal Education: Journal of Legal Education, 1952, 4, 441-447.

The criteria for item B would be fulfilled by almost all high school graduates in the top 10% of their class. Thus high ranking high school graduates who score very well on aptitude tests such as the SAT or LSAT have fulfilled the requirements of the committee. There is no real conflict; they should be admitted, especially since specific college course material is not required.

From the available evidence it seems that many students now being admitted to law schools throughout the nation do not have what the committee expects; the low scores on the law aptitude test give evidence of this. The law schools might be better advised to select top quality high school graduates rather than weak college students.

It is apparent from material published in the legal education literature that legal leaders are attempting to improve matters. I was particularly impressed by the following statement in the "Report of the Committee on Admissions to the Bar"[23] of which George Neff Stevens was chairman.

"Bar examination statistics covering the results of each examination should be made available showing the success of applicants according to pre-legal education, type of law school or other legal training and other information of value to prospective students, members of the legal profession and to members of the public who are interested in standards for admission to the bar."

Further evidence of concern for standards is to be found in publications of the American Bar Association. A recent pamphlet,[24] for example, presented the names of 31 law schools not approved by the American Bar Association. Such listings indicate stimulation for self improvement.

---

[23] Association of American Law Schools. Program and Reports of Committees: 1958 Annual Meeting. Chicago: Foundation Press, and West Publishing, 1958.

[24] Section of Legal Education and Admissions to the Bar. Law Schools and Bar Admission Requirements in the United States: 1958 Review Of Legal Education. Chicago: American Bar Association, 1958.

That some changes are necessary can be seen from studies such as that by Wagner Thielens, Jr.[25] He presented data which he felt were typical on students entering an Eastern law school and an Eastern medical school (same institution?), He indicated that while 30% of the students at this medical school believe "A large fraction of the public is suspicious of doctors," 72% of the law entrants subscribe to a similar statement about lawyers. Thielens also asked the entrants how many felt their career choice was the only one that could really satisfy them. The percentages: Medical, 64; Law 30. When they were asked how many felt they had "a great deal" of competition in their pre-professional college work the percentages were: Medical, 49; Law, 18. Of the medical entrants about 30% had A records, of the law entrants, 10%.

My sixth and final suggestion is to institute national bar examinations as one of the best means of controlling standards and affecting the public's image of lawyers. Since, as Dean Neff Stevens pointed out during an excellent symposium,[26] there are "about 10 or 12 subjects that over 90 per cent of the state boards of bar examiners listed for examination" it would seem that an objectively scoreable examination could be prepared for administration to all candidates in the country (say twice each year). Those who pass this examination which would be prepared, administered, and scored by some single agency would then be eligible to write an essay examination prepared, administered, and scored by the state boards. Those who pass the national and state examinations would qualify for a general law certificate

---

[25] Thielens, Wagner Jr. Some Comparisons of Entrants to Medical and Law School. Journal of Legal Education, 1958, 11, 153-170.

[26] Clark, Herbert W., Eliott, Sheldon, Stevens, George N,, and others. Nationally Administered Bar Examinations. Journal of Legal Education, 1954, 7, 28-48.

(or license?) From the legal literature it is apparent that only a minority of leaders (essay writers!) have confidence in examinations that do not require written answers. There appears to be abundant evidence that almost anything that can be measured by means of an essay exam can be measured better by means of an objective one. Everett S. Elwood[27] of the National Board of Medical Examiners explained why they replaced essay exams by objective ones: This change was not made by the National Board until it had spent nearly two years in testing the multiple choice type of examination and comparing the results with those obtained from the essay type of question by the same candidate on the same day. Results of both types of examination were correlated with the relative standing of the same examinees in their classes in medical school. The correlation was much higher between school standing and the essay type examination grades. The Board, therefore, felt fully justified in making the change to the type of examination that has broader scope and the greater validity.

One might say that "testing in law" is different from "testing in medicine" but I am confident the results of a systematic study would be the same. Even ability to write has been shown[28] to be measured and predicted more accurately by means of objective tests. One of the main reasons the validity of essay examinations is lower lies in the "reliability" of grading: several independent judgments of the written

---

[27] Clark, Herbert W., and Others. Op. cit. 26.

[28] Pearson, Richard. Should the General Composition Test be Continued? The Test Fails as an Entrance Examination. College Board Review, 1955, 25, 2-9 and

Huddleston, Edith M. Measurement of Writing Ability at the College Entrance Level: Objective Vs Subjective Testing Techniques. Ph.D. Thesis. New York University, 1952.

material do not agree, so all cannot be right. The essay and the selection interview have much in common. It is mostly because I am certain that too many have an unshakeable faith in the essay that I suggest it be retained in the bar examinations.

I believe the standard on the objective bar examination should be set fairly low. A fairly low standard would probably be high enough to eliminate almost all weak graduates and candidates from unapproved schools, but would still assure each state a substantial number to be examined by means of the local essay examination. States desiring to maintain very high standards might fail a large number by means of the essay exam; states having a shortage of lawyers would be in a position to pass all who have met the national standard. It is unfortunate that this would be necessary but reality must be faced.

A second set of examinations should also be available to but not required of those who pass both the national and state general examinations. The second set of examinations would be specialty examinations, also prepared nationally and locally. The standards for passing these examinations should be quite high, perhaps high enough so that only about one quarter of those who qualify for general certificates would obtain specialty certificates.

Each candidate would be required to pass objective examinations in, say, three fields from a possible 15 and then pass a state essay examination in one other field. The 15 fields in which objective specialty exams would be available might be torts, contracts, criminal law, real prpperty, personal property, agency, corporations, constitutional law, wills, negotiable instruments, equity, evidence, pleading, conflict of laws, and partnerships. These are the ones Dean Neff[29] indicated as being fairly

---

[29] Clark, Herbert W., and others. Op. cit. 26.

common to all states. The essay exam would be in some other field (e.g., trade regulation, domestic relations, and taxation.) Those passing the three national and one local exam should be given specialty certificates which they should be encouraged to display in appropriate ways. Perhaps it would be wise to permit a candidate to take no more than one set of specialty exams each year. The requirement of passing an essay exam again allows each state to control somewhat the number of lawyers with specialty certificates. National standards and the public interest would be served by having the objective examination as the initial screen.

The cost to candidates would probably be less than at present since successive hurdles are involved, and not all candidates would reach the expensive essay examinations. Possibly the general objective examination might be about $20. The general essay examination (which should be about one quarter the length of present bar exams) might be about $40. The specialty exams would cost more since they would be taken by a relatively small number.

High status and complicated and interesting problems would come to those in possession of specialty certificates while the general certificates would provide a "quality floor" for law. Standards could gradually be raised and quality carefully controlled.

The need seems great for some concrete improvements in the present bar assessment approach. Only one of the critical comments made by Eugene Glenn,[30] former President of the National Conference of Bar Examiners, is repeated here. He said, "It is an amazing thing that only three or four states throughout the forty nine admitting jurisdictions ever purport to make any statistical studies of the results to

---

[30] Clark, Herbert W., and others. Op. cit. 26.

ascertain the correlation" between bar examination performance and other criteria I too am impressed not only with the lack of "bar exam" oriented research but also with the lack of personnel and assessment studies of all kinds within the law profession The situation with regard to research on medical school applicants is in marked contrast For example, a 228 page issue of the Journal of Medical Education was recently devoted to "The Appraisal of Applicants to Medical Schools," a report of a four day conference on the topic.

The representatives of a small number of law schools deserve great credit for having launched, with the cooperation of the Educational Testing Service, the Law School Admission Test, a test of aptitude for the law. Legal educators and bar officials with the assistance of the Educational Testing Service or some other agency might well take another big step by launching the general and specialty Law Certification Tests, tests of proficiency in the law.

In conclusion it should be emphasized that it is far from necessary to accept the details of the suggestions presented in order for benefit to be derived from the underlying theoretical structure. The spirit rather than the letter of the proposals is of prime importance. It is my sincere hope that this paper, though dealing with admittedly controversial matters, will prove of value to you.

For Meeting in Ann Arbor, Michigan
June, 1959